

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-163096

(43)Date of publication of application : 16.06.2000

(51)Int.Cl. G10L 19/12
G10L 19/08
H03M 7/30
H04B 14/04

(21)Application number : 10-337805

(71)Applicant : NEC CORP

(22)Date of filing : 27.11.1998

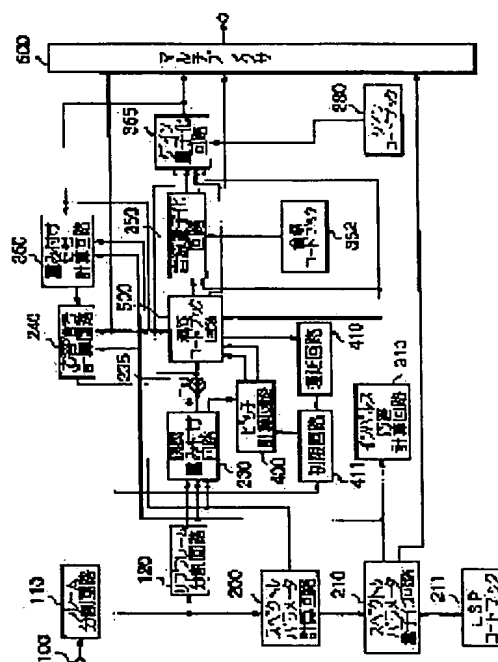
(72)Inventor : ITO HIRONORI
OZAWA KAZUNORI
SERIZAWA MASAHIRO

(54) SPEECH CODING METHOD AND SPEECH CODING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a speech coding device capable of obtaining excellent sound quality even at a lower bit rate in a CELP(Code Excited Linear Predictive coding).

SOLUTION: A pitch calculation circuit 400 calculates a pitch cycle T_{op} based on a searching range outputted from a limit circuit 411 by using an output signal $X_w(n)$ of an acoustic feeling weighting circuit 230, and outputs at least one pitch cycle to an adaptive code book circuit 500. The limit circuit 411 limits the pitch cycle searching range by using a delay of the adaptive code book requested by a former sub-frame outputted from a delay circuit 410.



LEGAL STATUS

[Date of request for examination] 27.11.1998

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3180786

[Date of registration] 20.04.2001

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-163096

(P2000-163096A)

(43) 公開日 平成12年6月16日 (2000.6.16)

(51) Int.Cl. ⁷	識別記号	F I	キーワード* (参考)
G 1 0 L 19/12		G 1 0 L 9/14	S 5 D 0 4 5
19/08			G 5 J 0 6 4
H 0 3 M 7/30		H 0 3 M 7/30	B 5 K 0 4 1
H 0 4 B 14/04		H 0 4 B 14/04	Z

審査請求 有 請求項の数 6 O L (全 11 頁)

(21) 出願番号 特願平10-337805

(22) 出願日 平成10年11月27日 (1998. 11. 27)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 伊藤 博紀

東京都港区芝五丁目7番1号 日本電気株式会社内

(72) 発明者 小澤 一範

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100100893

弁理士 渡辺 勝 (外3名)

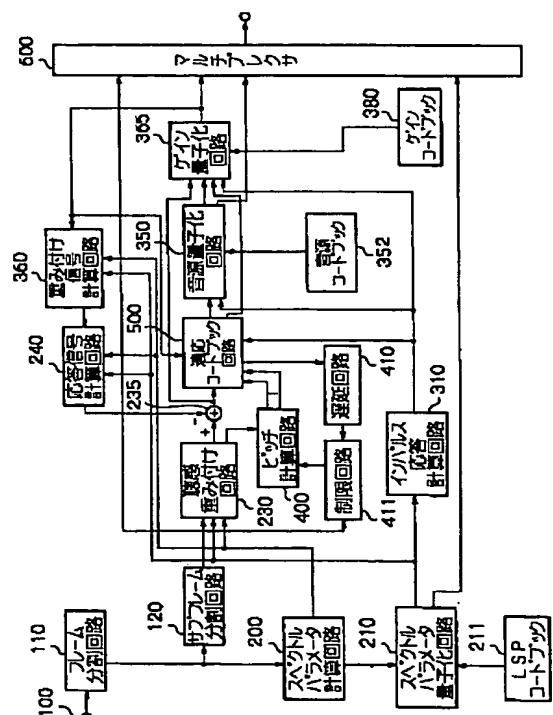
最終頁に続く

(54) 【発明の名称】 音声符号化方法及び音声符号化装置

(57) 【要約】

【課題】 CELP (Code Excited Lenear Predictive coding) において、より低ビットレートでも良好な音質の得られる音声符号化装置を提供する。

【解決手段】 ピッチ計算回路400において、聴感重み付け回路230の出力信号 $X_w(n)$ を用いて、制限回路411から出力される探索範囲に基づきピッチ周期 T_{op} を計算し、適応コードブック回路500に少なくとも1つのピッチ周期を出力する。制限回路411では、遅延回路410から出力される前サブフレームで求められた適応コードブックの遅延を用いて、ピッチ周期探索範囲を制限する。



【特許請求の範囲】

【請求項1】 入力された音声信号からスペクトルパラメータを求めて量子化する工程と、
ピッチ周期に基づき、過去に量子化された音源信号から適応コードブックにより遅延とゲインを求めるとともに音声信号を予測して残差を求める工程と、
前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化する工程と、
前記音源信号のゲインを量子化する工程と、
過去に求めた適応コードブックの遅延をもとに前記ピッチ周期の探索の際の探索範囲を制限して、前記音声信号から前記ピッチ周期を探索する工程と、を有する音声符号化方法。

【請求項2】 入力された音声信号からスペクトルパラメータを求めて量子化する工程と、
ピッチ周期に基づき、過去に量子化された音源信号から適応コードブックにより遅延とゲインを求めるとともに音声信号を予測して残差を求める工程と、
前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化する工程と、
前記音源信号のゲインを量子化する工程と、
前記音声信号から特徴量を抽出してモードを判別する工程と、
判別されたモードがあらかじめ定められたモードの場合には過去に求めた適応コードブックの遅延をもとに前記ピッチ周期の探索の際の探索範囲を制限して、前記音声信号から前記ピッチ周期を探索する工程と、を有する音声符号化方法。

【請求項3】 音声信号を入力しスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部と、
前記音声信号からピッチ周期を求め出力するピッチ計算部と、
前記ピッチ計算部の出力をもとに過去に量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求め遅延とゲインを出力する適応コードブック部と、
前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化し出力する音源量子化部と、
前記音源信号のゲインを量子化し出力するゲイン量子化部と、
ピッチ周期を探索する際に、過去に求めた適応コードブックの遅延をもとに探索範囲を制限する制限部とを有し、
前記制限部の出力をもとに前記ピッチ計算部がピッチ周期を探索し出力する音声符号化装置。

【請求項4】 音声信号を入力しスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部と、
前記音声信号からピッチ周期を求め出力するピッチ計算

部と、

前記ピッチ計算部の出力をもとに過去に量子化された音源信号から適応コードブックにより複数の遅延とゲインを求め音声信号を予測して残差を求め遅延とゲインを出力する適応コードブック部と、
前記スペクトルパラメータを用いて前記複数の遅延の各々に対し前記音声信号の音源信号を量子化し歪の小さい方を選択し出力する音源量子化部と、
前記音源信号のゲインを量子化し出力するゲイン量子化部と、
ピッチ周期を探索する際に、過去に求めた適応コードブックの遅延をもとに探索範囲を制限する制限部とを有し、
前記制限部の出力をもとに前記ピッチ計算部がピッチ周期を探索し出力する音声符号化装置。

【請求項5】 音声信号を入力しスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部と、
前記音声信号からピッチ周期を求め出力するピッチ計算部と、
前記ピッチ計算部の出力をもとに過去に量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求め遅延とゲインを出力する適応コードブック部と、
前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化し出力する音源量子化部と、
前記音声信号から特徴量を抽出してモードを判別し出力する判別部と、
前記音源信号のゲインを量子化し出力するゲイン量子化部と、
前期判別部の出力があらかじめ定められたモードの場合に、過去に求めた適応コードブックの遅延をもとに、ピッチ周期を探索する際の探索範囲を制限する制限部と、を有し、
前記あらかじめ定められたモードの場合に、前記ピッチ計算部が、前記制限部の出力をもとにピッチ周期を探索し出力する音声符号化装置。

【請求項6】 音声信号を入力しスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部と、
前記音声信号からピッチ周期を求め出力するピッチ計算部と、
前記ピッチ計算部の出力をもとに過去に量子化された音源信号から適応コードブックにより複数の遅延とゲインを求め音声信号を予測して残差を求め遅延とゲインを出力する適応コードブック部と、
前記スペクトルパラメータを用いて前記複数の遅延の各々に対し前記音声信号の音源信号を量子化して歪の小さい方を選択し出力する音源量子化部と、
前記音声信号から特徴量を抽出してモードを判別し出力

する判別部と、
前記音源信号のゲインを量子化し出力するゲイン量子化部と、
前記判別部の出力があらかじめ定められたモードの場合に、過去に求めた適応コードブックの遅延をもとに、ピッチ周期を探索する際の探索範囲を制限する制限部と、
を有し、
前記あらかじめ定められたモードの場合に、前記ピッチ計算部が、前記制限部の出力をもとにピッチ周期を探索し出力する音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声信号を低いビットレートで高品質に符号化するための音声符号化方法及び音声符号化装置に関するものである。

【0002】

【従来の技術】音声信号を高効率に符号化する方式としては、例えば、M. Schroeder and B. Atal氏による"Code-excited linear prediction: High quality speech at very low bit rates" (Proc. ICASSP, pp.937-940, 1985年)と題した論文(文献1)や、Kleijn氏らによる"Improved speech quality and efficient vector quantization in SELP" (Proc. ICASSP, pp.155-158, 1988年)と題した論文(文献2)などに記載されているCELP (Code Excited Linear Predictive Coding) が知られている。

【0003】CELPでは、送信側では、フレーム毎(例えば20ms)に音声信号から線形予測(LPC: Linear Predictive Coding)分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。フレームをさらにサブフレーム(例えば5ms)に分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータ(ピッチ周期に対応する遅延パラメータとゲインパラメータ)を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。ピッチ予測して求めた音源信号に対して、予め定められた種類の雑音信号からなる音源コードブック(ベクトル量子化コードブック)から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化するように行う。そして、選択されたコードベクトルの種類を表すインデックスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせる。

【0004】

【発明が解決しようとする課題】しかしながら、上述したCELPでは、現サブフレームで抽出される適応コードブックの遅延が、前サブフレームで計算された適応コ

ードブックの遅延に対し、整数倍以上あるいは整数の逆数倍以下となった場合(ここで整数は2以上のある整数)、前サブフレームと現サブフレームで適応コードブックの遅延が不連続となり、音質が劣化するという問題点があった。その理由は、現サブフレームで抽出される適応コードブックの遅延は、ピッチ計算部により音声信号から計算したピッチ周期付近で探索されるが、このピッチ周期が前サブフレームで計算された適応コードブックの遅延に対し、整数倍以上あるいは整数の逆数倍以下となった場合、現サブフレームに対する適応コードブックの探索範囲が、前サブフレームの適応コードブックの遅延付近を含まないので、前サブフレームと現サブフレームで適応コードブックの遅延が時間的に不連続になることである。

【0005】本発明の目的は、上述の問題を解決し、音質の劣化の少ない音声符号化方法及び装置を提供することにある。

【0006】

【課題を解決するための手段】本発明の音声符号化方法は、入力された音声信号からスペクトルパラメータを求めて量子化する工程と、ピッチ周期に基づき、過去に量子化された音源信号から適応コードブックにより遅延とゲインを求めるとともに音声信号を予測して残差を求める工程と、スペクトルパラメータを用いて音声信号の音源信号を量子化する工程と、音源信号のゲインを量子化する工程と、過去に求めた適応コードブックの遅延をもとにピッチ周期の探索の際の探索範囲を制限して、音声信号からピッチ周期を探索する工程と、を有する。

【0007】本発明の音声符号化方法では、音声信号から特徴量を抽出してモードを判別する工程をさらに設け、判別されたモードがあらかじめ定められたモードの場合にのみピッチ周期の探索の際の探索範囲が制限されるようにしてもよい。

【0008】本発明の音声符号化装置は、音声信号を入力しスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部と、音声信号からピッチ周期を求め出力するピッチ計算部と、ピッチ計算部の出力をもとに過去に量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求め遅延とゲインを出力する適応コードブック部と、スペクトルパラメータを用いて音声信号の音源信号を量子化し出力する音源量子化部と、音源信号のゲインを量子化し出力するゲイン量子化部と、ピッチ周期を探索する際に、過去に求めた適応コードブックの遅延をもとに探索範囲を制限する制限部とを有し、制限部の出力をもとにピッチ計算部がピッチ周期を探索し出力する。

【0009】本発明の音声符号化装置では、適応コードブック部が複数の遅延とゲインを求めるようにし、音源量子化部が、複数の遅延の各々に対し音声信号の音源信号を量子化し歪みの小さい方を選択し出力するようにして

もよい。さらに、音声信号から特徴量を抽出してモードを判別し出力する判別部を設け、あらかじめ定められたモードの場合に、ピッチ周期を探索する際の探索範囲が制限されるようにしてもよい。

【0010】(作用) 本発明では、制限部で、前サブフレームで求めた適応コードブックの遅延を入力し、前サブフレームで求めた適応コードブックの遅延と現在求める適応コードブックの遅延が不連続にならないようにピッチ周期探索範囲を制限し、そのピッチ周期探索範囲をピッチ計算部に出力する。ピッチ計算部では、聴感重み付け出力信号と制限部から出力されるピッチ周期探索範囲を入力し、ピッチ周期を計算し、適応コードブック部に少なくとも1つのピッチ周期を出力する。適応コードブック部では、聴感重み付け信号とゲイン量子化部から出力される過去の音源信号とインパルス応答計算回路から出力される聴感重み付けインパルス応答とピッチ計算部から出力されるピッチ周期を入力し、このピッチ周期近傍を探索して、適応コードブックの遅延を求める。これらの構成により、サブフレーム毎に求められる適応コードブックの遅延が時間的に不連続になることを避けることができる。

【0011】

【発明の実施の形態】次に、本発明の好ましい実施の形態について、図面を参照して説明する。

【0012】《第1の実施形態》図1は、本発明の第1の実施形態の音声符号化装置の構成を示すブロック図である。この音声符号化装置は、本件出願人による特開平8-320700号公報(文献3)に示す音声符号化装置とほぼ同様の構成の音声符号化装置に、ピッチ計算回路400、遅延回路410及び制限回路411を付加した構成のものである(なお、特開平8-320800号公報に記載のものでは、ゲインコードブックが2組設けられているが、ここでは、ゲインコードブックは1つ設けられている)。

【0013】この音声符号化装置では、入力端子100から入力した音声信号をフレーム(例えば20ms)ごと分割するフレーム分割回路110が設けられ、フレームは、サブフレーム分割回路120とスペクトルパラ

$$D_j = \sum_{i=1}^{10} W(i) [LSP(i) - QLSP(i)_j]^2 \quad (1)$$

ここで、 $LSP(i)$ 、 $QLSP(i)_j$ 、 $W(i)$ は、それぞれ、量子化前の*i*次目のLSP、量子化後の*j*番目の結果及び重み係数である。

【0017】以下では、量子化法として、ベクトル量子化を用いるものとし、第4サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。具体的な方法は例えば、特開平4-171500号公報(文献6)、特開平4-363000号公報(文献7)、特開平5-6199号公報(文献8)や、T. Nom

メータ計算回路200に出力される。サブフレーム分割回路120は、フレームの音声信号をフレームよりも短いサブフレーム(例えば5ms)に分割する。

【0014】スペクトルパラメータ計算回路200は、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓(例えば24ms)をかけて音声信号を切り出し、スペクトルパラメータをあらかじめ定められた次数(例えばP=10次)で計算する。ここでスペクトルパラメータの計算には、周知のLPC分析や、Burg分析等を用いることができる。ここでは、Burg分析を用いることとする。Burg分析の詳細については、中溝著による“信号解析とシステム同定”と題した単行本(コロナ社1988年刊)の82~87頁(文献4)等に記載されているので、その説明は省略する。さらにスペクトルパラメータ計算回路200では、Burg法により計算された線形予測係数 α_i ($i=1, \dots, 10$)が量子化や補間に適したLSP(線スペクトル対: line spectrum pair)パラメータに変換される。ここで、線形予測係数からLSPへの変換は、菅村他による“線スペクトル対(LSP)音声分析合成方式による音声情報圧縮”と題した論文(電子通信学会論文誌, J64-A, pp.599-606, 1981年)(文献5)を参照することができる。例えば、第2, 4サブフレームでBurg法により求めた線形予測係数をLSPパラメータに変換し、それによって第1, 3サブフレームのLSPを直線補間により求め、補間によって求めたLSPを逆変換して線形予測係数に戻し、第1~4サブフレームの線形予測係数 α_{il} ($i=1, \dots, 10$, $l=1, \dots, 5$)を聴感重み付け回路230に出力する。また、第4サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0015】スペクトルパラメータ量子化回路210は、LSPコードブック211を参照することにより、あらかじめ定められたサブフレームのLSPパラメータを効率的に量子化し、下式の歪み D_j を最小化する量子化値を出力する。

【0016】

【数1】

ura氏らによる“LSP Coding Using VQ-SVQ With Interpolation in 4.075 kbps M-LCELP Speech Coder”と題した論文(Proc. Mobile Multimedia Communications, pp. B.2.5, 1993)(文献9)等を参照できるので、ここでは説明は省略する。

【0018】また、スペクトルパラメータ量子化回路210は、第4サブフレームで量子化したLSPパラメータをもとに、第1~第4サブフレームのLSPパラメータを復元する。ここでは、現フレームの第4サブフレームの量子化LSPパラメータと1つ過去のフレームの第

4サブフレームの量子化LSPパラメータとを用いて直線補間を行い、(現フレームの)第1～第3サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後に、直線補間により第1～第4サブフレームのLSPを復元することできる。さらに性能を向上させるためには、前記誤差電力を最小化するコードベクトルを複数候補選択したのちに、各々の候補について、累積歪を評価し、累積歪を最小化する候補と補間LSPの組を選択するようにすることができる。詳細は、例えば、特開平6-222797号公報(文献10)を参照すればよい。

【0019】スペクトルパラメータ量子化回路210は、以上により復元した第1～3サブフレームのLSPと第4サブフレームの量子化LSPをサブフレームごとに線形予測係数 α'_{il} ($i=1, \dots, 10, l=1, \dots, 5$)に変換し、インパルス応答計算回路310へ出力する。また、第4サブフレームの量子化LSPのコードベクトルを表すインデックスをマルチプレクサ600に出力する。

【0020】スペクトルパラメータ計算回路200、ス

$$g_{op} = \frac{\sum_{n=1}^L X_w(n) \times X_w(n+T_{op})}{\sum_{n=1}^L \sqrt{X_w^2(n+T_{op})}} \quad (2)$$

ここで、Lはピッチ分析長を示す。ここでピッチ計算回路400は、音声信号からピッチ周期を求めて出力するピッチ計算部であり、制限回路411は、ピッチ周期を探索する際に、過去に求めた適応コードブックの遅延をもとに探索範囲を制限する制限部である。

【0024】遅延回路410が適応コードブック回路500と制限回路411との間に設けられており、この遅延回路410は、適応コードブック回路500から現サブフレームの適応コードブックの遅延を入力し、その値を次のサブフレームの処理まで保存し、1つ前のサブフ

区間1	17, 18, 19, 20, ..., 31, 32, 33, 35
区間2	36, 37, 38, 39, ..., 68, 69, 70, 71
区間3	72, 73, 74, 75, ..., 141, 142, 143, 144

例えば、前サブフレームで求められた適応コードブックの遅延が区間1に属しているならば、探索区間は、区間1と区間2に制限される。ここで、ピッチ周期探索範囲の分割テーブルとして表1以外のものを用いること、あるいは、テーブルを時間的に変更させることも可能である。

【0028】応答信号を計算するための応答信号計算回路240は、スペクトルパラメータ計算回路200から各サブフレームごとに線形予測係数 α_{il} を入力し、スベ

クトルパラメータ量子化回路210及びLSPコードブック211によって、入力音声信号のスペクトルパラメータを求めて量子化し出力するスペクトルパラメータ計算部が構成されている。

【0021】また、この音声符号化装置には、聴感重み付けを行うための聴感重み付け回路230が設けられている。聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に量子化前の線形予測係数 α_{il} ($i=1, \dots, 10, l=1, \dots, 5$)を入力し、前記文献1にもとづき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号 $X_w(n)$ を出力する。

【0022】ピッチ計算回路400は、聴感重み付け回路230の出力信号 $X_w(n)$ と制限回路411から出力されるピッチ周期探索範囲とを入力とし、このピッチ周期探索範囲内でピッチ周期 T_{op} を計算し、適応コードブック回路500に少なくとも1つのピッチ周期を出力する。ピッチ周期 T_{op} は、前記探索範囲において、以下の式を最大化するものが選択される。

【0023】

【数2】

レームの適応コードブックの遅延を制限回路411に出力する。

【0025】制限回路411は、遅延回路410から出力される、前サブフレームで求めた適応コードブックの遅延を入力とし、ピッチ周期探索範囲を出力する。この制限は、例えば、以下のように行う。

【0026】まず、探索するピッチ周期の範囲を表1で示すような3つの区間に分割したテーブルを用意する。

【0027】

【表1】

区間1	17, 18, 19, 20, ..., 31, 32, 33, 35
区間2	36, 37, 38, 39, ..., 68, 69, 70, 71
区間3	72, 73, 74, 75, ..., 141, 142, 143, 144

クトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α'_{il} をサブフレームごとに入力し、保存されているフィルタメモリの値を用いて、入力信号を零 $[d(n)=0]$ とした応答信号を1サブフレーム分計算し、減算器235へ出力する。ここで、応答信号 $x_z(n)$ は下式で表される。

【0029】

【数3】

$$x_z(n) = d(n) - \sum_{i=1}^{10} \alpha_i d(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i y(n-i) + \sum_{i=1}^{10} \alpha_i' \gamma^i x_z(n-i) \quad (3)$$

ただし、 $n-i \leq 0$ のときは

$$y(n-i) = p(N + (n-i)) \quad (4)$$

$$x_z(n-i) = s_w(N + (n-i)) \quad (5)$$

ここで N はサブフレーム長を示す。 γ は聴感重み付け量を制御する重み係数であり、下記の式(8)と同一の値である。 $s_w(n)$ 、 $p(n)$ は、それぞれ、重み付け信号計算回路360の出力信号、後述の式(7)における右辺第1項のフィルタの分母の項の出力信号をそれぞれ示す。重み付け信号計算回路360については後述する。

$$D_T = \sum_{n=0}^{N-1} x_w'^2(n) - \left[\sum_{n=0}^{N-1} x_w(n) y_w(n-T) \right]^2 / \left[\sum_{n=0}^{N-1} y_w^2(n-T) \right] \quad (6)$$

さらに、量子化されたスペクトルパラメータからインパルス応答を計算するインパルス応答計算回路310が設けられている。インパルス応答計算回路310は、 z 変換が下式で表される聴感重み付けフィルタのインパルス応答 $h_w(n)$ をあらかじめ定められた点数 L だけ計算

$$H_w(z) = \frac{1 - \sum_{i=1}^{10} \alpha_i z^{-i}}{1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}} \quad (7)$$

適応コードブック回路500は、ピッチ計算回路400の出力をもとに、過去に量子化された音源信号から適応コードブックをにより遅延 T とゲイン β を求め、音声信号を予測して残差(予測残差信号 $e_w(n)$)を求め、遅延 T 、ゲイン β 及び予測残差信号 $e_w(n)$ を出力するものである。適応コードブック回路500には、後述するゲイン量子化回路365から過去の音源信号 $v(n)$ が、減算器235から出力信号 $x'_w(n)$ が、インパルス応答計算回路310から聴感重み付けインパルス応答 h

$$D_T = \sum_{n=0}^{N-1} x_w'^2(n) - \left[\sum_{n=0}^{N-1} x'_w(n) y_w(n-T) \right]^2 / \left[\sum_{n=0}^{N-1} y_w^2(n-T) \right]^2 \quad (8)$$

ただし、

$$y_w(n-T) = v(n-T) * h_w(n) \quad (9)$$

式(9)において、記号*は畳み込み演算を表す。次に、適応コードブック回路500は、ゲイン β を下式に従い求める。

【0034】

【数7】

$$\beta = \sum_{n=0}^{N-1} x_w(n) y_w(n-T) / \sum_{n=0}^{N-1} y_w^2(n-T) \quad (10)$$

ここで、女性音や、子供の声に対して、適応コードブックの遅延の抽出精度を向上させるために、適応コードブックの遅延を整数サンプルではなく、小数サンプル値で求めてもよい。具体的な方法は、例えば、P.Kroon氏らによる、"Pitchpre-dictors with high temporal resolution"と題した論文(Proc. ICASSP, pp.661-664, 1990

【0030】減算器235は、下式により、聴感重み付け回路230が出力する聴感重み付け信号 $X_w(n)$ から応答信号 $x_z(n)$ を1サブフレーム分減算し、 $x'_w(n)$ を適応コードブック回路500へ出力する。

【0031】

【数4】

し、適応コードブック回路500及び音源量子化回路350へ出力する。

【0032】

【数5】

$w(n)$ が、ピッチ計算回路400からピッチ周期 T_{op} がそれぞれ入力する。そして適応コードブック回路500は、このピッチ周期 T_{op} 近傍を探索して、下式の歪みを最小化するように適応コードブックの遅延 T を求め、適応コードブックの遅延を表すインデックスをマルチプレクサ600に出力する。さらに、適応コードブックの遅延の値を遅延回路410にも出力する。

【0033】

【数6】

年)(文献11)等を参照することができる。

【0035】さらに、適応コードブック回路500は、式(10)に従いピッチ予測を行ない、予測残差信号 $e_w(n)$ を音源量子化回路350へ出力する。

【0036】

【数8】

$$e_w(n) = x_w(n) - \beta v(n-T) * h_w(n) \quad (11)$$

音源量子化回路350は、スペクトルパラメータを用いて音声信号の音源信号を量子化し出力するものであって、音源信号として M 個のパルスをたてる。また、音源量子化回路350は、パルスの振幅を M パルス分まとめて量子化するための、 B ビットの振幅コードブック、もしくは極性コードブックを有している。以下では、極性

コードブックを用いる場合の説明を行なう。この極性コードブックは、音源コードブック352に格納されている。

【0037】音源量子化回路350は、音源コードブック352に格納された各極性コードベクトルを読み出し、各コードベクトルに対して位置をあてはめ、式(12)を最小化するコードベクトルと位置の組合せを複数セット選択する。

【0038】

【数9】

$$\Phi(n) = \sum_{i=0}^{N-1} e_w(i) h_w(i-n), n=0, \dots, N-1 \quad (12)$$

ここで、 $h_w(n)$ は、聴感重み付けインパルス応答である。式(12)を最小化するには、式(13)を最大化する極性コードベクトル g_{ik} と位置 m_i の組合せを求めればよい。

【0039】

【数10】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} e_w(n) s_{wk}(m_i) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m_i) \quad (13)$$

パルス番号	位置
第1パルス	0, 5, 10, 15, 20, 25, 30, 35
第2パルス	1, 6, 11, 16, 21, 26, 31, 36
第3パルス	2, 7, 12, 17, 22, 27, 32, 37
第4パルス	3, 8, 13, 18, 23, 28, 33, 38
第5パルス	4, 9, 14, 19, 24, 29, 34, 39

極性コードベクトルの探索終了後、音源量子化回路350は、選択された複数セットの極性コードベクトルと位置の組み合わせをゲイン量子化回路365に出力する。

【0042】ゲイン量子化回路365は、音源信号のゲインを量子化して出力するものであり、音源量子化回路350から、複数セットの極性コードベクトルとパルス位置の組み合わせが入力する。ゲイン量子化回路365は、ゲインコードブック380からゲインコードベクトル

$$D_k = \sum_{n=0}^{N-1} \left[x_w(n) - \beta_i v(n-T) * h_w(n) - G'_i \sum_{i=1}^M g'_{ik} h_w(n-m_i) \right]^2 \quad (16)$$

ここでは、ゲイン量子化回路365において、適応コードブックのゲインとパルスで表した音源のゲインの両者を同時にベクトル量子化する例について示した。ゲイン量子化回路365は、ゲイン選択された極性コードベクトルを表すインデックス、位置を表す符号、ゲインコードベクトルを表すインデックスをマルチプレクサ600に出力する。

【0044】なお、複数パルスの振幅を量子化するためのコードブックを、音声信号を用いてあらかじめ学習して格納しておくこともできる。コードブックの学習法は、例えば、Linde氏らによる"An algorithm for vecto

または、式(14)を最大化するように選択してもよい。この方が、式中の分子の計算に要する演算量が低減化される。

【0040】

【数11】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} \Phi(n) v_k(n) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m_i) \quad (14)$$

ただし、

$$\Phi(n) = \sum_{i=0}^{N-1} e_w(i) h_w(i-n), n=0, \dots, N-1 \quad (15)$$

ここで、各パルスのとり得る位置は、演算量削減のため、文献4に示すように、拘束することができる。一例として、 $N=40$ 、 $M=5$ とすると、各パルスのとり得る位置は表2のようになる。

【0041】

【表2】

ルを読み出し、選択された複数セットの極性コードベクトルと位置の組み合わせに対して、式(16)を最小化するようにゲインコードベクトルを探索し、歪みを最小化するゲインコードベクトル、極性コードベクトルと位置の組み合わせを1種類選択する。

【0043】

【数12】

r quantization design," と題した論文(IEEE Trans. Commun., pp.84-95, January, 1980) (文献12) 等を参照できる。

【0045】ここで重み付け信号計算回路360について説明する。重み付け信号計算回路360は、それぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読み出した後、式(17)にもとづき駆動音源信号 $v(n)$ を求める。

【0046】

【数13】

$$v(n) = \beta'_i v(n-T) * h_w(n) + G'_i \sum_{i=1}^M g'_i \delta(n-m_i) \quad (17)$$

この駆動音源信号 $v(n)$ は、適応コードブック回路 500 に出力される。次に、重み付け信号計算回路 360 は、スペクトルパラメータ計算回路 200 の出力パラメータ、スペクトルパラメータ量子化回路 210 の出力パラメータを用いて式(18)により、応答信号 $s_w(n)$ をサ

$$s_w(n) = v(n) - \sum_{i=1}^{10} \alpha_i v(n-i) + \sum_{i=1}^{10} \alpha_i \tau^i p(n-i) + \sum_{i=1}^{10} \alpha_i \tau^i s_w(n-i) \quad (18)$$

マルチプレクサ 600 には、スペクトルパラメータ量子化回路 210 から第 4 サブフレームの量子化 LSP のコードベクトルを表すインデックスが入力し、音源量子化回路 350 から極性コードベクトルと位置との組み合わせが入力し、ゲイン量子化回路 365 から極性コードベクトルを表すインデックス、位置を表す符号、ゲインコードベクトルを表すインデックスが入力する。マルチプレクサ 600 は、これらの入力に基づいて、サブフレームに分割された音声信号に対する符号を組み立てて出力する。以上により、入力音声信号の音声符号化がなされたことになる。

【0048】この音声符号化装置では、制限回路 411 が、前サブフレームで求めた適応コードブックの遅延を入力として、前サブフレームで求めた適応コードブックの遅延と現在求める適応コードブックの遅延が不連続にならないようにピッチ周期探索範囲を制限し、そのピッチ周期探索範囲をピッチ計算回路 400 に出力する。ピッチ計算回路 400 は、聴感重み付け回路 230 の出力信号 $X_w(n)$ と制限回路 411 から出力されるピッチ周期探索範囲を入力として、ピッチ周期 T_{op} を計算し、適応コードブック回路 500 に少なくとも 1 つのピッチ周期 T_{op} を出力する。適応コードブック回路 500 は、聴感重み付け信号 $x'_w(n)$ と、ゲイン量子化回路 365 から出力される過去の音源信号 $v(n)$ と、インパルス応答計算回路 310 から出力される聴感重み付けインパルス応答 $h_w(n)$ と、ピッチ計算回路 400 から出力されるピッチ周期 T_{op} とを入力として、このピッチ周期近傍を探索して、適応コードブックの遅延を求める。これらの構成により、サブフレーム毎に求められる適応コードブックの遅延が時間的に不連続になることを避けることができる。

【0049】《第 2 の実施形態》次に、本発明の第 2 の実施形態の音声符号化装置について、図 2 を参照して説明する。この音声符号化装置は、図 1 に示す装置と比較して、適応コードブック回路と音源量子化回路の動作が異なっている。図 2 において図 1 と同一の番号を付したものは、図 1 に示したものと同一の構成要素である。

【0050】適応コードブック回路 511 は、適応コードブックの遅延を式(8)を最小化するようにして求め、複数の候補を音源量子化回路 351 に出力する。これらの候補に対して、音源量子化回路 351 とゲイン量子化

ブフレーム毎に計算し、応答信号計算回路 240 へ出力する。

【0047】

【数 14】

回路 365 において、第 1 の実施形態と同様に音源とゲインの量子化が行なわれ、最終的にすべての候補の中で、式(16)を最小化する組み合わせが 1 種類選択される。他の動作は、第 1 の実施形態の場合と同様である。

【0051】この音声符号化装置においても、過去に求めた適応コードブックの遅延に基づいてピッチ周期の探索範囲が制限されるので、サブフレーム毎に求められる適応コードブックの遅延が時間的に不連続になることを避けることができる。

【0052】《第 3 の実施形態》次に、本発明の第 3 の実施形態の音声符号化装置について、図 3 を参照して説明する。この音声符号化装置は、図 1 に示す装置と比較して、モード判定回路 800 が付加されており、また、制限回路の動作が異なっている。図 3 において図 1 と同一の番号を付したものは、図 1 に示したものと同一の構成要素である。モード判別回路 800 を設けて複数のモードを設定することにより、ここでは図示していないがモードに応じて適応コードブック回路 500 の動作条件を変えることなどが可能になり、モードごとに最適な符号化を設定できるようになって、低ビットレートでより高品質な音声符号化を行うことができる。

【0053】モード判別回路 800 は、聴感重み付け回路 230 の出力信号を用いて、特徴量を抽出し、フレーム毎にモードの判別を行う。ここで、特徴量としては、ピッチ予測ゲインを用いることができる。サブフレーム毎に求めたピッチ予測ゲインをフレーム全体で平均し、この値とあらかじめ定められた複数のしきい値を比較し、あらかじめ定められた複数のモードに分類する。ここでは、一例として、モードの種類が 4 種類であるとする。この場合、モード 0, 1, 2, 3 は、それぞれ、無声区間、過渡区間、弱い有声区間、強い有声区間にほぼ対応するものとする。このモードに従って、制限回路 412 は、例えば、モード 0 では、ピッチ周期探索範囲の制限を行わず、モード 1, 2, 3 では、ピッチ周期探索範囲の制限を行うというような探索範囲の切り替えを行う。なお、判別されたモードを示す情報は、モード判別回路 800 からマルチプレクサ 600 にも出力されている。他の動作は、第 1 の実施形態の場合と同様である。

【0054】《第 4 の実施形態》次に、本発明の第 4 の実施形態の音声符号化装置について、図 2 を参照して説明する。この音声符号化装置は、図 2 に示す装置と比較し

て、モード判定回路800が付加されており、また、制限回路の動作が異なっている。図4において図2と同一の番号を付したものは、図2に示したものと同一の構成要素である。第3の実施形態と同様にモード判別回路800を設けて複数のモードを設定することにより、低ビットレートでより高品質な音声符号化を行うことができる。

【0055】モード判別回路800は、聴感重み付け回路230の出力信号を用いて、特徴量を抽出し、フレーム毎にモードの判別を行う。ここで、特徴量としては、ピッチ予測ゲインを用いることができる。サブフレーム毎に求めたピッチ予測ゲインをフレーム全体で平均し、この値とあらかじめ定められた複数のしきい値を比較し、あらかじめ定められた複数のモードに分類する。ここでは、一例として、モードの種類が4種類であるとする。この場合、モード0、1、2、3は、それぞれ、無声区間、過渡区間、弱い有声区間、強い有声区間にほぼ対応するものとする。このモードに従って、制限回路412は、例えば、モード0では、ピッチ周期探索範囲の制限を行わず、モード1、2、3では、ピッチ周期探索範囲の制限を行うというような探索範囲の切り替えを行う。なお、判別されたモードを示す情報は、モード判別回路800からマルチプレクサ600にも出力されている。他の動作は、第2の実施形態の場合と同様である。

【0056】

【発明の効果】以上説明したように本発明は、前サブフレームの適応コードブックの遅延に基づいてピッチ計算部でのピッチ周期探索範囲を制限することにより、サブフレームごとに求められる適応コードブックの遅延が時間的に不連続になることが避けられ、従来よりも良好な音質で符号化できるようになる、という効果がある。

【図面の簡単な説明】

【図1】本発明の第1の実施形態の音声符号化装置の構成を示すブロック図である。

【図2】本発明の第2の実施形態の音声符号化装置の構成を示すブロック図である。

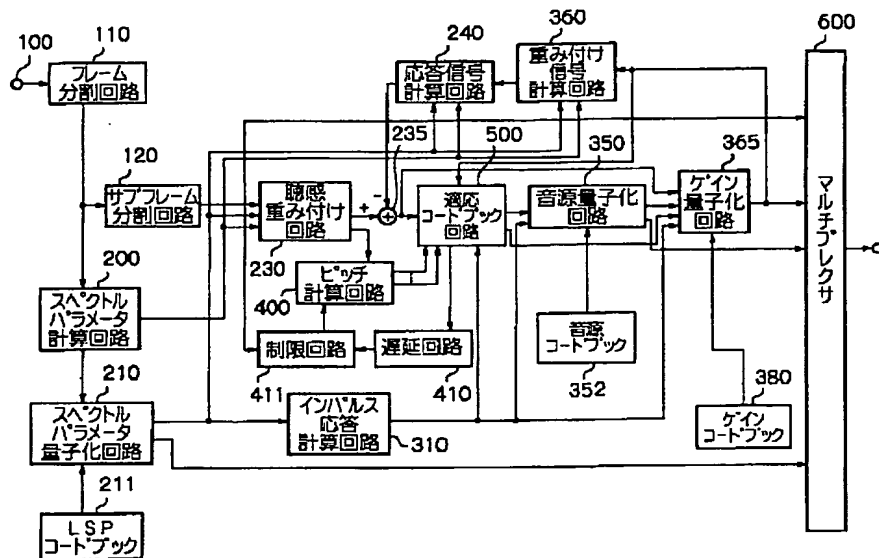
【図3】本発明の第3の実施形態の音声符号化装置の構成を示すブロック図である。

【図4】本発明の第4の実施形態の音声符号化装置の構成を示すブロック図である。

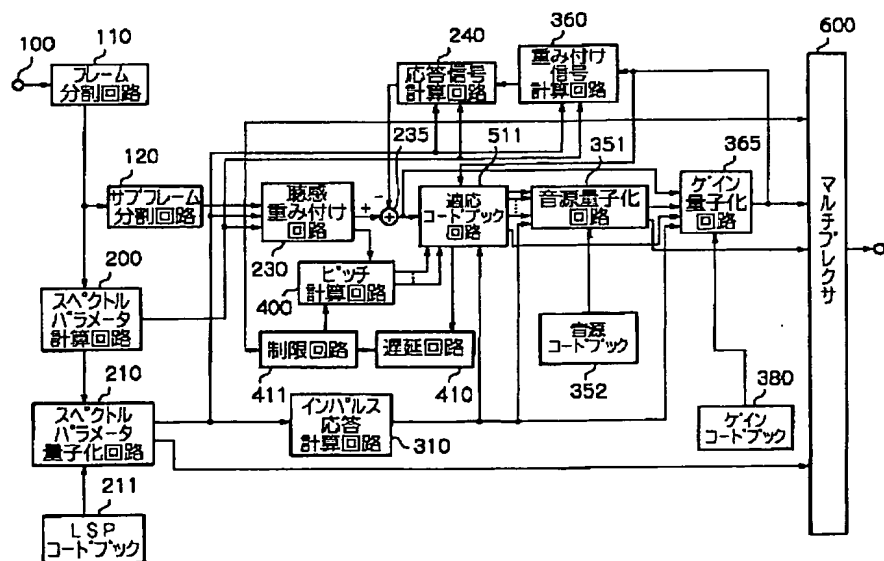
【符号の説明】

100	入力端子
110	フレーム分割回路
120	サブフレーム分割回路
200	スペクトルパラメータ計算回路
210	スペクトルパラメータ量子化回路
211	LSPコードブック
230	聴感重み付け回路
235	減算器
240	応答信号計算回路
310	インパルス応答計算回路
350, 351	音源量子化回路
352	音源コードブック
360	重み付け信号計算回路
365	ゲイン量子化回路
380	ゲインコードブック
410	遅延回路
411, 412	制限回路
500, 511	適応コードブック回路
600	マルチプレクサ
800	モード判別回路

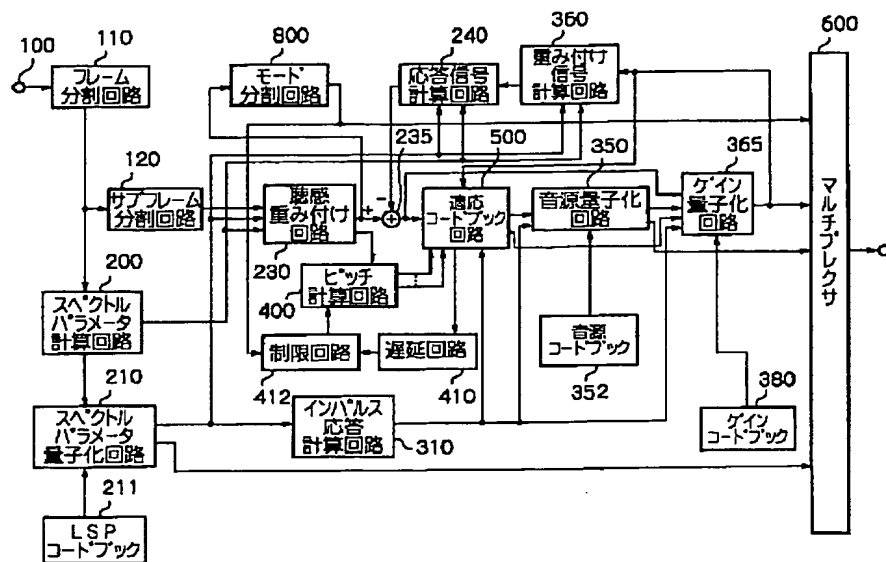
【図1】



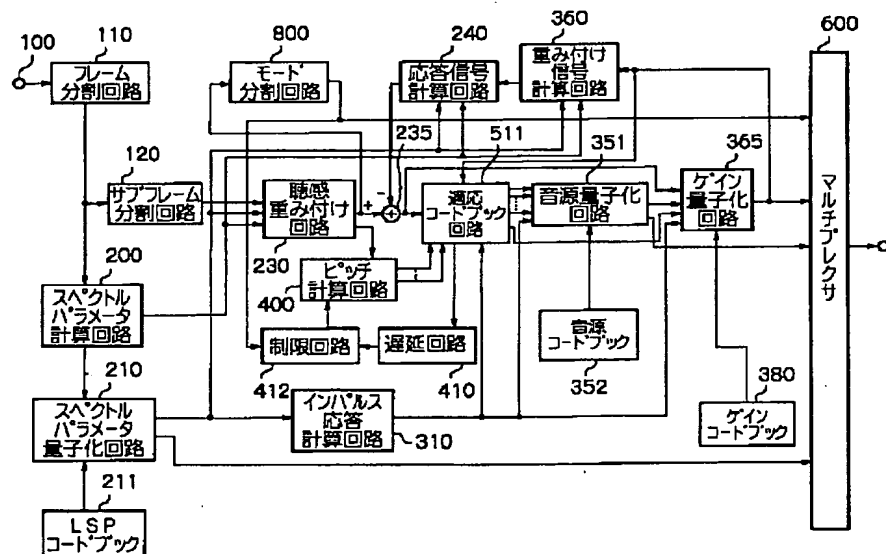
【図2】



【図3】



【図4】



フロントページの続き

(72)発明者 芹沢 昌宏
東京都港区芝五丁目7番1号 日本電気株
式会社内

Fターム(参考) 5D045 CA01 DA11
5J064 AA01 BA13 BC01 BC02 BC26
BD02
5K041 AA07 BB02 CC01 DD09 EE14
EE35 FF31 FF40 HH22 HH25
HH37 HH40